|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | |  | **INTRODUCTION** | |  |  |  |
|  | |  | 1.1 Overview | |  |  |  |
|  | |  | 1.2 Purpose | |  |  |  |
| **2** | |  | **LITERATURE SURVEY** | |  |  |  |
|  | |  | 2.1 Existing problem | | |  |  |
|  | |  | 2.2 Proposed solution | | |  |  |
| **3** | |  | **THEORITICAL ANALYSIS** | |  |  |  |
|  | |  | 3.1 Block diagram | |  |  |  |
|  | |  | 3.2 Hardware / Software designing | | |  |  |
| **4** | |  | **EXPERIMENTAL INVESTIGATIONS** | | |  |  |
| **5** | |  | **FLOWCHART** | |  |  |  |
| **6** | |  | **RESULT** |  |  |  |  |
| **7** | |  | **ADVANTAGES & DISADVANTAGES** | | |  |  |
| **8** | |  | **APPLICATIONS** | |  |  |  |
| **9** | |  | **CONCLUSION** | |  |  |  |
| **10** | |  | **FUTURE SCOPE** | | **`** |  |  |
| **11** | |  | **BIBILOGRAPHY** | |  |  |  |
|  | |  | **APPENDIX** | |  |  |  |
|  |  | | A. Source code | |  |  |  |
|  |  | |  |  |  |  |  |

# INTRODUCTION

## Overview:

This is a typical Regression Machine Learning project which leverages on historical data obtained from Kaggle website to predict insights into the future. This problem statement is aimed at predicting Life Expectancy rate of a country given various features.

Life expectancy is the factor which show a formidable way to give the expectation value of a person years to live by analysed historical records of preceding years. This factor can contribute in determining other factor for survival of human beings, one of these factors can be the supply and demand flow which ultimately support the economy and keep up the flow.

Thus, this project is about the prediction of the life expectation of a person which is the expected average lifetime of beings in the world.

## Purpose:

The prediction of life expectancy supports to keep up the flow in human society, by affecting fields from economy to research and development. The newest of studies by research can be obtained through the life expectancy value for the future, which can further contribute for the benefits of the society.

# LITERATURE SURVEY

## Existing Problem:

Need of a machine learning model for predicting the life expectancy.

Life expectancy is a statistical measure of the average time a human being is expected to live, life expectancy depends on various factors: Regional variations, Economic Circumstances, Sex Differences, Mental Illnesses, Physical Illnesses, Education, Year of their birth and other demographic factors. This project provides a way to predict average life expectancy of people living in a country when various factors such as year, GDP, education, alcohol intake of people in the country, expenditure on healthcare system and some specific disease related deaths that happened in the country are given.

## Proposed Solution:

The solution for predicting the life expectancy is to build a machine leaning model which can predict the future values by using the historical data. By investigating the data it is clear that the problem requires regression model as the predicting value is a continuous value.

The model can work to predict values by providing input which can be implemented by the using an user-interface for input-output.

# THEORITICAL ANALYSIS

## Block Diagram:

## Hardware / Software designing:

The model for machine learning consists of python notebook with the implemented algorithms and the web UI flow. The model is designed based on a Regression algorithm.

The required steps to build the model are as follows:

* Getting the dataset on life expectancy by WHO.
* Analyze the factors affecting the life expectancy.
* Data preprocessing or cleaning the dataset.
* Splitting the data and train it.
* Predict the value.
* Debugging the code.
* Deploy the Model.

The services used for this task are:

IBM Cloud with services such as Machine Learning, Watson Studio and Node RED.

Here, Node Red is used as the mean to create a web user interface.

# EXPERIMENTAL INVESTIGATIONS

The investigation shows that the given problem is a regression problem as the values for the life expectancy are continuous and not discrete.

Therefore, implementing a regression algorithm is the optimal option to get predictions as Regression model methods allow to predict a continuous outcome variable (y) based on the value of one or multiple predictor variables (x). Briefly, the goal of **regression model** is to build a mathematical equation that defines y as a function of the x variables.

Here in this project, Linear Regression algorithm is applied to predict values. Linear Regression algorithm is based on supervised learning and here it facilitates to find the relationship between variables (like GDP, population, status, etc.) and forecasting (life expectancy) for the model to be trained to give out predictions.

# FLOWCHART



# RESULT

The result of the project gives out prediction of life expectancy value with:

r2\_score as accuracy: 80% approx. (79.9709 exactly)

Mean Absolute Error: 3.205027483377878

Mean Squared Error: 17.45972465686501

Root Mean Squared Error: 4.178483535550309

|  |  |  |  |
| --- | --- | --- | --- |
| ADVANTAGES & DISADVANTAGES **The advantages of the project are as follows:**   * Gave a better perspective on how to analyze given dataset and find relationship among different variables in dataset * Helped us in learning how to make machine learning models in python and using it to predict future values for a given variable. * Taught us about IBM cloud and its services and how to integrate it with our machine learning model. * Taught me how to make a Node-Red application and integrate it with our machine learning model.   **The disadvantages of the project are as follows:**   * The project just gives a brief introduction about IBM Cloud * The project doesn’t teach us about how to improvise our model | |  | |
| APPLICATIONS The skills acquired from this project can be used:   * To make various IBM cloud services and execute python code in IBM Watson Studio. * To build a Node-RED application, integrate it with other services and deploy it online. * To make machine learning models and integrate it with IBM Cloud services. |  | |
| CONCLUSIONThrough this project I was able to gain a holistic view on IBM cloud and its services such as machine learning services, Node-RED services and IBM Watson Studio. This project also gave me an insight on machine learning with Python and various steps that we have to go through in order to build an accurate machine learning model and deploy it in IBM Cloud. Thus, I was able to predict the life expectancy of a person using the dataset that was provided and various machine learning services. This trained me for further practical applications of machine learning in real life situations and prepared me for future opportunities in machine learning. |  | |
| FUTURE SCOPESince I was able to make a machine learning model for predicting the life expectancy of a person using python and IBM cloud, I can now help communities and fields such as the medical industry, business field, etc. This will allow me to contribute to social welfare and people’s well-being. I will also be provided with various job opportunities because of the skills I acquired from this project. |  | |
| BIBLIOGRAPHY Kaggle (Dataset):  <https://www.kaggle.com/kumarajarshi/life-expectancy-who>  IBM Cloud:  <https://cloud.ibm.com/> |  | |
| APPENDIX |  | |
| A. Source code Github repository (containing source code):  Predicting Model Notebook:  <https://github.com/SmartPracticeschool/llSPS-INT-1582-Predicting-Life-Expectancy-using-Machine-Learning/blob/master/Predicting%20Life%20expectancy%20using%20Machine%20Learning.ipynb> |  | |

Node Red flow:

[https://github.com/SmartPracticeschool/llSPS-INT-1582-Predicting-Life-Expectancy- using-Machine-Learning/blob/master/flows.json](https://github.com/SmartPracticeschool/llSPS-INT-1582-Predicting-Life-Expectancy-%20%20using-Machine-Learning/blob/master/flows.json)

Dataset(used in the model ):

<https://github.com/SmartPracticeschool/llSPS-INT-1582-Predicting-Life-Expectancy-using-Machine-Learning/blob/master/Life%20Expectancy%20Data.csv>